

10/573,262

## Sequence Alignments

Q68HV1\_MOUSE

ID Q68HV1\_MOUSE PRELIMINARY; PRT; 1925 AA.  
 AC Q68HV1;  
 DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.  
 DT 11-OCT-2004, sequence version 1.  
 DT 13-JUN-2006, entry version 11.  
 DE Plexin D1.  
 GN Name=Plxnd1;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muroidea; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=BALB/c;  
 RA Duke-Cohan J.S., Ahmed W., Reinherz E.L.;  
 RT "Identification of mouse plexin D1."  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; AY688678; AAT99561.1; -; mRNA.  
 DR UniGene; Mm.3085; -.  
 DR Ensembl; ENSMUSG00000030123; Mus musculus.  
 DR MGI; MGI:2154244; Plxnd1.  
 DR RZPD-ProtExp; IOM17766; -.  
 DR GO; GO:0005886; C:plasma membrane; IC.  
 DR GO; GO:0005515; F:protein binding; IPI.  
 DR GO; GO:0017154; F:semaphorin receptor activity; IPI.  
 DR GO; GO:0001569; P:patterning of blood vessels; IMP.  
 DR InterPro; IPR002909; IPT\_TIG\_rcpt.  
 DR InterPro; IPR003659; Plexin-like.  
 DR InterPro; IPR013548; Plexin\_cytopl.  
 DR InterPro; IPR002165; Plexin\_repeat.  
 DR InterPro; IPR008936; Rho\_GAP.  
 DR InterPro; IPR001627; Sema.  
 DR Pfam; PF08337; Plexin\_cytopl; 1.  
 DR Pfam; PF01437; PSI; 2.  
 DR Pfam; PF01403; Sema; 1.  
 DR Pfam; PF01833; TIG; 3.  
 DR SMART; SM00429; IPT; 3.  
 DR SMART; SM00423; PSI; 3.  
 DR SMART; SM00630; Sema; 1.  
 DR PROSITE; PS51004; SEMA; 1.  
 SQ SEQUENCE 1925 AA; 211638 MW; 53E71ADBD05A78FC CRC64;

Query Match 100.0%; Score 9181; DB 2; Length 1925;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SMLNVAANHPNASTVGLVLPPTSGTGGSRLLVGATYTGFGSAFFPRNRSLEDHRFENTPE 60

Db	180	 SMLNVAANHPNASTVGLVLPPTSGTGGSRLLVGATYTGFGSAFFPRNRSLEDHRFENTPE	239
Qy	61	IAIRSLDARGDLAKLFTFDLNPSTGGSRLLVGATYTGFGSAFFPRNRSLEDHRFENTPE	120
Db	240	IAIRSLDARGDLAKLFTFDLNPSTGGSRLLVGATYTGFGSAFFPRNRSLEDHRFENTPE	299
Qy	121	YLALNSEARAGDKDSQARSLLARICLPRGAGGDAKKLTESYIQLGLQCAGGAGRGDLYSR	180
Db	300	YLALNSEARAGDKDSQARSLLARICLPRGAGGDAKKLTESYIQLGLQCAGGAGRGDLYSR	359
Qy	181	LVSVPFAPAREQFFAVFERPQGAPGARNAPAALCAFRFDDVQAAIRAARTACFVEPAPDVVA	240
Db	360	LVSVPFAPAREQFFAVFERPQGAPGARNAPAALCAFRFDDVQAAIRAARTACFVEPAPDVVA	419
Qy	241	VLDSVVQGTGPACESKRNIQLQPEQLDCGAAHLQHPLTILQPLRASPVFRAPGLTAVAVA	300
Db	420	VLDSVVQGTGPACESKRNIQLQPEQLDCGAAHLQHPLTILQPLRASPVFRAPGLTAVAVA	479
Qy	301	SANNYTAVFLGTATGRLLKISLNESMQVVSRRVLTVAAYGEPVHHVMQFDPMDPGYLYLMT	360
Db	480	SANNYTAVFLGTATGRLLKISLNESMQVVSRRVLTVAAYGEPVHHVMQFDPMDPGYLYLMT	539
Qy	361	SHQMARVKVAACEVHSTCGDCVGAADAYCGWCTLETRCTLQQDCTNSSQPHFWTSASEGP	420
Db	540	SHQMARVKVAACEVHSTCGDCVGAADAYCGWCTLETRCTLQQDCTNSSQPHFWTSASEGP	599
Qy	421	SRCPAMTVLPSEIDVHRDYGTMILQISGSLPSLSGEMACDYGNVVRTVARVPGPAYDHQ	480
Db	600	SRCPAMTVLPSEIDVHRDYGTMILQISGSLPSLSGEMACDYGNVVRTVARVPGPAYDHQ	659
Qy	481	IAYCNLLPRAQFSPFAGQDHVTVEMSVRVKGNIVSANFTIYDCSRIGQVYPHTACTSC	540
Db	660	IAYCNLLPRAQFSPFAGQDHVTVEMSVRVKGNIVSANFTIYDCSRIGQVYPHTACTSC	719
Qy	541	LSTQWPCSWCIQLHSCVSNQSQCDSPNPTSPQDCPQILPSPLAPVPTGGSQDILVPLTK	600
Db	720	LSTQWPCSWCIQLHSCVSNQSQCDSPNPTSPQDCPQILPSPLAPVPTGGSQDILVPLTK	779
Qy	601	ATFFHGSSLECSFGLSEESFEAVWANNLSLVRNQNQVVLHTTQKSQVFPLSLKLKGPDPDRFLD	660
Db	780	ATFFHGSSLECSFGLSEESFEAVWANNLSLVRNQNQVVLHTTQKSQVFPLSLKLKGPDPDRFLD	839
Qy	661	SPNPMTVVVYNCAMGSPDCSQCLGREDLGHLGVWWDGCRRLRGPLQPLPGTCPAPEIRAIE	720
Db	840	SPNPMTVVVYNCAMGSPDCSQCLGREDLGHLGVWWDGCRRLRGPLQPLPGTCPAPEIRAIE	899
Qy	721	PLSGPLDGGTLLTIRGRNLGRRLSDVAHGVIWIGSVACEPLADRYTVSEEIVCATGPAAGA	780
Db	900	PLSGPLDGGTLLTIRGRNLGRRLSDVAHGVIWIGSVACEPLADRYTVSEEIVCATGPAAGA	959
Qy	781	FSDVVTVNVSKGRSREQFSYVLPTVHSLPSMGPKAGGTRITIHGSDLVNVSMLQVLVN	840
Db	960	FSDVVTVNVSKGRSREQFSYVLPTVHSLPSMGPKAGGTRITIHGSDLVNVSMLQVLVN	1019
Qy	841	DTDPTDLTRTATSITCTVPGGTLPSVPVVCVRFESRGCVHGNLTFWYMQNPVITAISPG	900

Db	1020	DTDPCDLTRTATSITCTVPGGTLPSVPVPCVRFESRGCVHGNTLFWYMQNPVITAISPG	1079
Qy	901	RSPVSGGRTITVAGERFHMVQNVSMVHHIGREPTFCKVLNSTLITCSPGALSNASAPV	960
Db	1080	RSPVSGGRTITVAGERFHMVQNVSMVHHIGREPTFCKVLNSTLITCSPGALSNASAPV	1139
Qy	961	DDFFINGRAYADEAAEELDPAEAQRGSRFRLDYLPNPQFSTAKREKWIKHHPGEPLTLVI	1020
Db	1140	DDFFINGRAYADEAAEELDPAEAQRGSRFRLDYLPNPQFSTAKREKWIKHHPGEPLTLVI	1199
Qy	1021	HKEQDSLGLSHEYHIKIGQVSCDIQIISDRVIHCSVNESLGTAEQQLPITIQVGNFNQT	1080
Db	1200	HKEQDSLGLSHEYHIKIGQVSCDIQIISDRVIHCSVNESLGTAEQQLPITIQVGNFNQT	1259
Qy	1081	IATLQLGGSETAIVVSIVICSVLLLLSVVALFVFCTKSRRRAERYWQKTLQMEEMESQIR	1140
Db	1260	IATLQLGGSETAIVVSIVICSVLLLLSVVALFVFCTKSRRRAERYWQKTLQMEEMESQIR	1319
Qy	1141	EEIRKGFAELQTDMDLTKELNRSQGIPFLEYKHFVTRTFFPKCSSLYEERYVLPSKTLN	1200
Db	1320	EEIRKGFAELQTDMDLTKELNRSQGIPFLEYKHFVTRTFFPKCSSLYEERYVLPSKTLN	1379
Qy	1201	SQGGSPQETHPLLGEWNIPEHCRPSMEEGISLFSLLNNKHFLIVFVHALEQQKDFAVR	1260
Db	1380	SQGGSPQETHPLLGEWNIPEHCRPSMEEGISLFSLLNNKHFLIVFVHALEQQKDFAVR	1439
Qy	1261	DRCSLASLLTIALHGKLEYYSIMKELLVDLIDASAAKNPKMLLRRTESVVEKMLTNWMS	1320
Db	1440	DRCSLASLLTIALHGKLEYYSIMKELLVDLIDASAAKNPKMLLRRTESVVEKMLTNWMS	1499
Qy	1321	ICMYGCLRETVGEPFFLLLCAIKQQINKGSIDAITGKARYTLNEEWLLRENIEAKPRNLN	1380
Db	1500	ICMYGCLRETVGEPFFLLLCAIKQQINKGSIDAITGKARYTLNEEWLLRENIEAKPRNLN	1559
Qy	1381	VSFQGCGMDSL SVRAMD TDTLTQVKEKILEAFCKNPYSQWPRAEDVDLEWFASSTQSYV	1440
Db	1560	VSFQGCGMDSL SVRAMD TDTLTQVKEKILEAFCKNPYSQWPRAEDVDLEWFASSTQSYV	1619
Qy	1441	LRDLDDTSVVEDGRKKLNTLAHYKIPEGASLAMS LTDKKDSTLGRVKDL DTEKYFHLVLP	1500
Db	1620	LRDLDDTSVVEDGRKKLNTLAHYKIPEGASLAMS LTDKKDSTLGRVKDL DTEKYFHLVLP	1679
Qy	1501	TDELVEPKKSHRQSHRKKVLPEIYLTRLLSTKGT LQKFLDDL FKA ILSIREDKPPLAVKY	1560
Db	1680	TDELVEPKKSHRQSHRKKVLPEIYLTRLLSTKGT LQKFLDDL FKA ILSIREDKPPLAVKY	1739
Qy	1561	FFDFLEEQA EKRGISDPDTLHIWKTNSLPLRFWVNILKNPQFVFDIEKTDHIDACLSVIA	1620
Db	1740	FFDFLEEQA EKRGISDPDTLHIWKTNSLPLRFWVNILKNPQFVFDIEKTDHIDACLSVIA	1799
Qy	1621	QAFIDACSI SLDQLGKDSPTNKLLYAKEIPEYRKTVQRYYKQIQDMTPLSEQEMNAHLAE	1680
Db	1800	QAFIDACSI SLDQLGKDSPTNKLLYAKEIPEYRKTVQRYYKQIQDMTPLSEQEMNAHLAE	1859
Qy	1681	ESRKYQNEFNTNVAMAEIYKYAKRYRPQIMAALEANPTARRTQLQHKFEQVVALMENNIY	1740
Db	1860	ESRKYQNEFNTNVAMAEIYKYAKRYRPQIMAALEANPTARRTQLQHKFEQVVALMENNIY	1919

Qy	1741	ECYSEA	1746
Db	1920	ECYSEA	1925

RESULT 4

PLXD1\_HUMAN

ID PLXD1\_HUMAN STANDARD; PRT; 1925 AA.  
AC Q9Y4D7; Q6PJS9; Q8IZJ2; Q9BTQ2;  
DT 31-AUG-2004, integrated into UniProtKB/Swiss-Prot.  
DT 31-AUG-2004, sequence version 2.  
DT 25-JUL-2006, entry version 36.  
DE Plexin-D1 precursor.  
GN Name=PLXND1; Synonyms=KIAA0620;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
OC Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).  
RC TISSUE=Brain;  
RX MEDLINE=98403880; PubMed=9734811; DOI=10.1093/dnares/5.3.169;  
RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.,  
RA Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. X.  
RT The complete sequences of 100 new cDNA clones from brain which can  
RT code for large proteins in vitro.";  
RL DNA Res. 5:169-176(1998).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), AND TISSUE SPECIFICITY.  
RX MEDLINE=22299888; PubMed=12412018; DOI=10.1002/dvdy.10159;  
RA van der Zwaag B., Hellemons A.J.C.G.M., Leenders W.P.J.,  
RA Burbach J.P.H., Brunner H.G., Padberg G.W., Van Bokhoven H.;  
RT "PLEXIN-D1, a novel plexin family member, is expressed in vascular  
RT endothelium and the central nervous system during mouse  
RT embryogenesis.";  
RL Dev. Dyn. 225:336-343(2002).  
RN [3]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 1386-1925 (ISOFORMS 1 AND  
RP 2).  
RC TISSUE=Muscle, and Uterus;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP GLYCOSYLATION [LARGE SCALE ANALYSIS] AT ASN-500, AND MASS  
 RP SPECTROMETRY.  
 RX PubMed=16335952; DOI=10.1021/pr0502065;  
 RA Liu T., Qian W.-J., Gritsenko M.A., Camp D.G. II, Monroe M.E.,  
 RA Moore R.J., Smith R.D.;  
 RT "Human plasma N-glycoproteome analysis by immunoaffinity subtraction,  
 RT hydrazide chemistry, and mass spectrometry."  
 RL J. Proteome Res. 4:2070-2080(2005).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q9Y4D7-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q9Y4D7-2; Sequence=VSP\_011516;  
 CC -!- TISSUE SPECIFICITY: Detected at low levels in heart, placenta,  
 CC lung, skeletal muscle, kidney, thymus and liver. Detected at very  
 CC low levels in brain, colon, spleen, small intestine and peripheral  
 CC blood leukocytes.  
 CC -!- SIMILARITY: Belongs to the plexin family.  
 CC -!- SIMILARITY: Contains 3 IPT/TIG domains.  
 CC -!- SIMILARITY: Contains 1 Sema domain.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; AB014520; BAA31595.1; ALT\_INIT; mRNA.  
 DR EMBL; AY116661; AAM49063.1; -; mRNA.  
 DR EMBL; BC003526; AAH03526.1; -; mRNA.  
 DR EMBL; BC011848; AAH11848.1; -; mRNA.  
 DR UniGene; Hs.301685; -.  
 DR Ensembl; ENSG00000004399; Homo sapiens.  
 DR HGNC; HGNC:9107; PLXND1.  
 DR MIM; 604282; gene.  
 DR InterPro; IPR002909; IPT\_TIG\_rcpt.  
 DR InterPro; IPR003659; Plexin-like.  
 DR InterPro; IPR013548; Plexin\_cytopl.  
 DR InterPro; IPR002165; Plexin\_repeat.  
 DR InterPro; IPR008936; Rho\_GAP.  
 DR InterPro; IPR001627; Sema.  
 DR Pfam; PF08337; Plexin\_cytopl; 1.  
 DR Pfam; PF01437; PSI; 2.  
 DR Pfam; PF01403; Sema; 1.  
 DR Pfam; PF01833; TIG; 3.  
 DR SMART; SM00429; IPT; 3.  
 DR SMART; SM00423; PSI; 3.  
 DR SMART; SM00630; Sema; 1.  
 DR PROSITE; PS51004; SEMA; 1.  
 KW Alternative splicing; Glycoprotein; Membrane; Polymorphism; Receptor;  
 KW Repeat; Signal; Transmembrane.  
 FT SIGNAL 1 46 Potential.  
 FT CHAIN 47 1925 Plexin-D1.  
 FT /FTId=PRO\_0000024676.  
 FT TOPO\_DOM 47 1271 Extracellular (Potential).  
 FT TRANSMEM 1272 1292 Potential.  
 FT TOPO\_DOM 1293 1925 Cytoplasmic (Potential).

FT	DOMAIN	47	546	Sema.
FT	DOMAIN	891	979	IPT/TIG 1.
FT	DOMAIN	981	1066	IPT/TIG 2.
FT	DOMAIN	1069	1160	IPT/TIG 3.
FT	CARBOHYD	86	86	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	155	155	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	188	188	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	224	224	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	481	481	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	500	500	N-linked (GlcNAc. . .).
FT	CARBOHYD	583	583	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	696	696	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	736	736	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	802	802	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	965	965	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1017	1017	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1060	1060	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1099	1099	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1118	1118	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1132	1132	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1237	1237	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1257	1257	N-linked (GlcNAc. . .) (Potential).
FT	VAR_SEQ	1766	1925	SLPLRFWVNILKNPQFVFDIDKTDHIDACLSVIAQAFIDAC
FT				SISDLQLGKDSPTNKLLYAKEIPEYRKIVQRYKQIQDMTP
FT				LSEQEMNAHLAEESRKYQNEFNTNVAMAEIYKYAKRYPQI
FT				MAALEANPTARRTQLQHKFEQVVALMEDNIYECYSEA ->
FT				RWRPSSPVLGEHPPEPPVCL (in isoform 2).
FT				/FTId=VSP_011516.
FT	VARIANT	870	870	M -> V (in dbSNP:2255703).
FT				/FTId=VAR_022144.
SQ	SEQUENCE	1925 AA;	212095 MW;	26001F5D0B2A80E5 CRC64;

Query Match 92.2%; Score 8468; DB 1; Length 1925;  
 Best Local Similarity 91.9%; Pred. No. 0;  
 Matches 1606; Conservative 52; Mismatches 88; Indels 2; Gaps 1;

Qy	1	SMLNVAANHPNASTVGLVLPPTSGTGGSRLLVGATYTGFGSAFFPRNRSLEDHRFENTPE	60
		:	
Db	178	SMLNVAANHPNASTVGLVLPAAAGAGGSRRLLVGATYTGYGSSFFPRNRSLEDHRFENTPE	237
Qy	61	IAIRSLDARGDLAKLFTFDLNPSSDDNLIKIKQGAKEQHKLGFVRAFLHPAVPPHSAQPYA	120
Db	238	IAIRSLDTRGDLAKLFTFDLNPSSDDNLIKIKQGAKEQHKLGFVSAFLHPSDPPPGAQSYA	297
Qy	121	YLALNSEARAGDKDSQARSLLARICLPRGAGGDAKKLTESYIQLGLQCAGGAGRGDLYSR	180
Db	298	YLALNSEARAGDKESQARSLLARICLPHGAGGDAKKLTESYIQLGLQCAGGAGRGDLYSR	357
Qy	181	LVSVPFAPAREQFFAVFERPQGAPGARNAPAALCAFRFDDVQAAIRAARTACFVEPAPDVVA	240
		:               :	
Db	358	LVSVPFAPARERLFVFERPQGSAPAARAAPAALCAFRFADVRAAIRAARTACFVEPAPDVVA	417
Qy	241	VLDSVVQGTGPACESKRNIQLQPEQLDCGAAHLQHPLTILQPLRASPVFRAPGLTAVAVA	300
		:           :	
Db	418	VLDSVVQGTGPACERKLNIQLQPEQLDCGAAHLQHPLSILQPLKATPVFRAPGLTSVAVA	477

Qy	301	SANNYTAVFLGTATGRLLKISLNESMQVVSRRVLT VAYGEPVHHVMQFDPMDPGYLYLMT	360
Db	478	SVNNYTAVFLGTVNGRLLKINLNESMQVVSRRVTVAYGEPVHHVMQFDPADSVYLYLMT	537
Qy	361	SHQMARVKVAACEVHSTCGDCVGAADAYCGWCTLETRCTLQQDCTNSSQPHFWTSASEGP	420
Db	538	SHQMARVKVAACNVHSTCGDCVGAADAYCGWCALETRCTLQQDCTNSSQQHFWTSASEGP	597
Qy	421	SRCPAMTVLPSEIDVHRDYGTMILQISGSLPSLSGMEMACDYGNGVRTVARVPGPAYDHQ	480
Db	598	SRCPAMTVLPSEIDVRQEYPGMILQISGSLPSLSGMEMACDYGNNIRTVARVPGPAFGHQ	657
Qy	481	IAYCNLLPRAQFPSFPAGQDHVTVEMSVRVKGNIVSANFTIYDCSRIGQVYPHTACTSC	540
Db	658	IAYCNLLPRDQFPFPNPDHVTVEMSVRVNGRNIVKANFTIYDCSRTAQVYPHTACTSC	717
Qy	541	LSTQWPCSWCIQLHSCVSNQSQCDSPNPTSPQDCPQILPSPLAPVPTGGSQDILVPLTK	600
Db	718	LSAQWPCFWCSQQHSCVSNQSRCEASNPTSPQDCPRTLSPAPVPTGGSQNILVPLAN	777
Qy	601	ATFFHGSSLECSFGLEESFEAVWANNLSVRCNQVVLHTTQKSQVFPLSLKLKGPDPDRFLD	660
Db	778	TAFFQGAALECSFGLEEIFEAVWVNESVVRCDQVVLHTTRKSQVFPLSLQLKGRPARFLD	837
Qy	661	SPNPMTVVVYNCAMGSPDCSQCLGREDLGHLGVWNDGCRRLRGPLQPLPGTCPAPEIRAIE	720
Db	838	SPEPMTVMVYNCAMGSPDCSQCLGREDLGHLGCMWSDGCRRLRGPLQPMAGTCPAPEIRAIE	897
Qy	721	PLSGPLDGGTLLTIRGRNLGRRLSDVAHGVSIGVACEPLADRYTVSEEIVCATGPAAGA	780
Db	898	PLSGPLDGGTLLTIRGRNLGRRLSDVAHGVSIGVACEPLPDYTVSEEIVCVTGPAAGP	957
Qy	781	FSDVVTVNVSKEGRSREQFSYVLPTVHSLPSMGPKAGGTRITIHGSDLNVGSMQLQVLVN	840
Db	958	LSGVVTVNASKEGKSRDRFSYVLPLVHSLPTMGPKAGGTRITIHGNDLHVGSSELQVLVN	1017
Qy	841	DTDPTDLTRTATSITCTVPGGTLPSVPVPCVRFESRGCVHGNTFWYMQNPVITAISPG	900
Db	1018	DTDPTCELMRTDTSIACMTPEGALPAPVPVPCVRFERRGCVHGNTFWYMQNPVITAISPR	1077
Qy	901	RSPVSGGRTITVAGERFHMVQNVSMVHHIGREPTFCKVLNSTLITCSPGALSNASAPV	960
Db	1078	RSPVSGGRTITVAGERFHMVQNVSMVHHIGREPTLCKVLNSTLITCSPGALSNASAPV	1137
Qy	961	DFFINGRAYADE--AAEELLDPAEAQRGSRFRLDYLPNPQFSTAKREKWIKHHPGEPLTL	1018
Db	1138	DFFINGRAYADEVAVAEELLDPEEAQRGSRFRLDYLPNPQFSTAKREKWIKHHPGEPLTL	1197
Qy	1019	VIHKEQDSLGLSHEYHIKIGQVSCDIQIISDRVIHCSVNESLGAEGQLPITIQVGNFN	1078
Db	1198	VIHKEQDSLGLQSHEYRVKIGQVSCDIQIVSDRIIHCSVNESLGAAGVQLPITIQVGNFN	1257
Qy	1079	QTIATLQLGGSETAIVSVIVICSVLLLLSVVALFVFCTKSRRARYWQKTLLQMEEMESQ	1138
Db	1258	QTIATLQLGGSETAIIIVSVIVICSVLLLLSVVALFVFCTKSRRARYWQKTLLQMEEMESQ	1317



Qy	1139	IREEIRKGFAELQTDMDLTKELNRSQGIPFLEYKHFVTRTFFPKCSSLYEERYVLPSKT	1198
Db	1318	IREEIRKGFAELQTDMDLTKELNRSQGIPFLEYKHFVTRTFFPKCSSLYEERYVLPSQT	1377
Qy	1199	LNSQGGSPQETHPLLGEWNIPEHCRPSMEEGISLFSSLLNNKHFLIVFVHALEQQKDFA	1258
Db	1378	LNSQGGSSQAQETHPLLGEWKIPESCRPNMEEGISLFSSLLNNKHFLIVFVHALEQQKDFA	1437
Qy	1259	VRDRCSLASLLTIALHGKLEYYSIMKELLVDLIDASAAKNPKMLLRRTESVVEKMLTNW	1318
Db	1438	VRDRCSLASLLTIALHGKLEYYSIMKELLVDLIDASAAKNPKMLLRRTESVVEKMLTNW	1497
Qy	1319	MSICMYGCLRETVGEPFFLLCAIKQQINKGSIDAITGKARYTLNEEWLLRENIEAKPRN	1378
Db	1498	MSICMYSCLRETVGEPFFLLCAIKQQINKGSIDAITGKARYTLNEEWLLRENIEAKPRN	1557
Qy	1379	LNVSFQGCMDLSLVRAMDTDTLTQVKEKILEAFCKNPYSQWPRAEDVDLEWFASSTQS	1438
Db	1558	LNVSFQGCMDLSLVRAMDTDTLTQVKEKILEAFCKNPYSQWPRAEDVDLEWFASSTQS	1617
Qy	1439	YVLRDLDDTSVVEDGRKKLNTLAHYKIPEGASLAMS LTDKKDSTLGRVKDL DTEKYFHLV	1498
Db	1618	YILRDLDDTSVVEDGRKKLNTLAHYKIPEGASLAMS LTDKKDNTLGRVKDL DTEKYFHLV	1677
Qy	1499	LPTDELVEPKKSHRQSHRKKVLPEIYLTRL LSTKGTLQKF LDDL FKA ILS IREDKPPLAV	1558
Db	1678	LPTDELAEPKKSHRQSHRKKVLPEIYLTRL LSTKGTLQKF LDDL FKA ILS IREDKPPLAV	1737
Qy	1559	KYFFDFLEEQAEKRGISDPDTLHIWKNTSLPLRFVWNILKNPQVFVDIEKTDHIDACLSV	1618
Db	1738	KYFFDFLEEQAEKRGISDPDTLHIWKNTSLPLRFVWNILKNPQVFVDIDKTDHIDACLSV	1797
Qy	1619	IAQAFIDACSIDLQLGKDSPTNKLLYAKEIPEYRKTVQRYYKQIQDMTPLSEQEMNAHL	1678
Db	1798	IAQAFIDACSIDLQLGKDSPTNKLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQEMNAHL	1857
Qy	1679	AEESRKYQNEFNTNVAMAEIYKYAKRYRPQIMAALEANPTARRTQLQHKFEQVVALMENN	1738
Db	1858	AEESRKYQNEFNTNVAMAEIYKYAKRYRPQIMAALEANPTARRTQLQHKFEQVVALMEDN	1917
Qy	1739	IYECYSEA	1746
Db	1918	IYECYSEA	1925